

## Table of Contents

<b>CHAPTER 8</b>	<b>DATA SOURCES OF NURSING HOME NURSE STAFFING ANALYSIS: ASSESSMENT OF OSCAR COMPARED TO MEDICAID COST REPORTS .....</b>	<b>8-1</b>
<b>8.1</b>	<b>Introduction .....</b>	<b>8-1</b>
<b>8.2</b>	<b>Data Sources .....</b>	<b>8-2</b>
8.2.1	Ohio Payroll Data .....	8-2
8.2.2	Online Survey and Certification Reporting System (OSCAR) Data .....	8-4
8.2.3	Medicaid Cost Report Data .....	8-5
8.2.4	Creation of Analytic Staffing Measure Variables .....	8-6
<b>8.3</b>	<b>Methods .....</b>	<b>8-6</b>
<b>8.4</b>	<b>Results .....</b>	<b>8-7</b>
8.4.1	Analysis of Average Staffing Levels .....	8-7
8.4.2	Correlation Analysis .....	8-9
8.4.3	Consistency of Identification of Low-Staffed Facilities .....	8-14
8.4.4	Conclusion of Validity Analyses: Medicaid Cost Report Data Are More Valid Than OSCAR .....	8-16
<b>8.5</b>	<b>Identifying Unreliable Staffing Data .....</b>	<b>8-17</b>
8.5.1	Identification of Extreme Outliers .....	8-17
8.5.2	Exclusions Based on Changes in Staffing Levels Across Time .....	8-18
<b>8.6</b>	<b>Potential Exclusion Criteria for Medicaid Cost Report Staffing Measures .....</b>	<b>8-21</b>
<b>8.7</b>	<b>Conclusions .....</b>	<b>8-23</b>

## **CHAPTER 8                      DATA SOURCES OF NURSING HOME NURSE STAFFING ANALYSIS: ASSESSMENT OF OSCAR COMPARED TO MEDICAID COST REPORTS<sup>1</sup>**

### **8.1      Introduction**

The purpose of this chapter is to compare the reliability and validity of staffing measures from the Online Survey Certification and Reporting System (OSCAR) to staffing measures derived from Medicaid Cost Report data. These analyses were used to determine the source of staffing data for the outcome analyses presented in Chapters 9 through 12 that analyze the relationship between staffing levels and resident outcomes.

Reliability and validity were tested based on comparison to a sample of nursing facilities in Ohio for which payroll records were collected. The payroll data provide close to a “gold standard” measure of facility nurse staffing. They were collected for the period corresponding to the most recently available OSCAR assessment and a second period, which typically covered the six months prior to OSCAR. This is the first data collection effort that captures both a similar definition and an identical time period as the OSCAR nurse staffing survey data using an independently collected and not self-reported facility data source. The validity analyses included a variety measures of correlation with the payroll data, including Pearson and Spearman correlation coefficients and Kendall’s Tau.

While the small sample size was a reason for caution (there were 78 facilities in the Ohio payroll sample for which both OSCAR and Cost Report data were available), our analyses suggested that the Cost Report data were more valid and reliable than OSCAR. This was especially true for lower staffed facilities, an important criterion given the interest in these facilities for the outcomes study. Because the OSCAR and payroll data covered the same time period, while the Cost Report data were based on the facility average across an entire year, the finding that staffing measures from the Cost Report data tended to be more highly correlated with the payroll data than staffing measures from OSCAR, strongly suggests that the Cost Report data were the more reliable and valid source of staffing data, and that

---

<sup>1</sup> This report was completed by Alan White of Abt Associates for the Health Care Financing Administration (Contract #500-95-0062-T.O.3; Allison Walker, Abt Associates Project Director; Marvin Feuerberg, HCFA Project Officer).

Other individuals who made valuable comments and suggestions on the analyses included in this chapter include Karen Reilly, Donna Hurd, and Terry Moore of Abt Associates, Andy Kramer and Michael Lin of the University of Colorado Health Center on Aging and Division of Geriatric Medicine, University of Colorado Health Sciences Center, Denver, Colorado, and Marvin Feuerberg, HCFA Project Officer. Mike Murphy of Abt Associates prepared the staffing files, which were based in part on OSCAR data for 1995-1997 prepared by Mick Cowles, of Cowles Research Group. Other valuable assistance was offered by Beth Klitch of Survey Solutions, Inc., who supervised the collection of Ohio payroll data, and Ed Mortimore and Sheila Lambowitz of the Health Care Financing Administration. Editorial assistance was provided by Susan Joslin, HCFA.

they should be the source of staffing measures for the outcomes study.

While the Cost Report data were more reliable than OSCAR, there were some facilities for which Cost Report data did not appear to be accurate. Two types of potential decision rules were developed to exclude facilities with questionable staffing measures. The first is a set of “logical” decision rules that identify facilities with obvious data errors. The second is based on the consistency of staffing measures across time and across data sources. While there are certainly some facilities that experience large changes in staffing levels across time, it is likely that a disproportionate share of data for these facilities are inaccurate. It may be desirable to exclude facilities that report large changes in staffing measures reported in the Cost Report data if these changes cannot be validated using OSCAR (or some other independent data source). A limitation of this type of exclusion criteria is that it relies on OSCAR to validate changes in Cost Report data, although it was found that the Cost Report data were more valid and reliable than OSCAR.

Decision rules that are based on *change* in reported staffing levels rather than the actual *level* of staffing (except in the case of extreme outliers) allows the distribution of staffing levels to be preserved, keeping low-staffed facilities with reliable data in the analysis. The payroll sample was not large enough to evaluate the impact of these decision rules on the validity of staffing measures for non-excluded facilities, but analysis of the impact of the decision rules on the consistency of staffing measures from the Cost Report and OSCAR data showed that application of the decision rules improved the consistency of staffing measures from the two data sources.

## **8.2 Data Sources**

Data sources used in this study included OSCAR, Medicaid Cost Reports, and Ohio payroll data. The payroll data were collected for a sample of 107 facilities, 78 of which also had Medicaid Cost Report data available. Medicaid Cost Report and OSCAR data for the years 1995-1997 were also available for other Ohio facilities and facilities from New York and Texas. A brief discussion follows. A detailed description of the Ohio payroll data collection can be found in Chapter 7.

### **8.2.1 Ohio Payroll Data**

Primary data collection was performed by Survey Solutions, Inc. Data collectors, experienced in skilled nursing facility administration, were trained to collect nursing staff payroll journal data closely following the OSCAR survey variable definitions (e.g., staff type, hours per week). The payroll data were logged using a computer laptop application. All nursing staff hours were captured, including full time, part time, and contract staff/organizations, as well as information regarding resident census. Other data sources, beyond payroll journals, were also used to capture staffing information. For example, staff time sheets, contract invoices, or other invoices were also considered potential sources of staffing documentation.

Total nursing staff payroll hours were collected for two time periods. The first reflects the two week time period prior to the most recent facility survey (i.e., the same time period used by the state survey agency). Payroll data were also collected for a second time period that reflected six months prior to the most recent facility survey. Where six month payroll data were unavailable, data were collected to encompass the oldest two week period available, provided there was at least one month between the first time period and the second.

HCFA's OSCAR data were used to capture records on Ohio facilities and stratify the facility into one of four nursing hours categories. Consideration was also given to ensure facility variation with respect to size, geographic distribution, for profit/not for profit status, and chain affiliation although these stratification requirements were less stringent. In all strata, except for the nursing hours per resident per day, the facility sample distribution generally parallels the Ohio facility distribution or facility average. Furthermore, the Ohio facility distribution is reflective of the US for these facility characteristics.

OSCAR was used to stratify Ohio facilities into one of four categories, based on total nursing hours per resident day. Consideration was also given to ensure facility variation with respect to size, geographic distribution, for profit/not for profit status, and chain affiliation, although these stratification requirements were less stringent. In all strata, except for the nursing hours per resident per day, the facility sample distribution generally parallels the Ohio facility distribution or facility average. Practical constraints limited the total sample to 107 facilities. Initially, 1997 OSCAR data were used to stratify facilities, but Abt gained access to the 1998 OSCAR data during the data collection process, and an updated facility category listing was generated using this more recent survey data. The goal was to select a stratified random sample based on facility staffing levels, using the following sample distribution.

<i>Category 1:</i>	Less than 2.0 total nursing hours per resident per day; 31 facilities;
<i>Category 2:</i>	2.0 - 2.5 total nursing hours per resident per day; 21 facilities;
<i>Category 3:</i>	2.6-3.6 total nursing hours per resident per day; 21 facilities;
<i>Category 4:</i>	Over 3.6 total nursing hours per resident per day; 34 facilities.

Low staffed facilities were over-represented in the sample design. In 1998, only 3% of Ohio facilities fell into Category One, 11% in Category 2, 59% in Category 3, and 27% in Category 4 (These figures are based on 1998 OSCAR data for Ohio facilities, before the application of any of the decision rules discussed later in this chapter).

Low-staffed facilities were over sampled in order to evaluate the reliability of OSCAR among low-staffed facilities. As a result, the payroll sample is not representative of Ohio facilities. Ohio had few facilities in Categories 1 or 2, and, based on staffing levels, the payroll sample was actually more representative of nursing homes nationwide than it is of Ohio facilities. Nationwide, nearly 8% of facilities had fewer than two nursing hours per resident day, while 17% had between 2.0 and 2.5 total hours, 53% used between 2.5 and 3.6 total hours, and 22% 3.6 or more total nursing hours per

resident day.

Due to an insufficient number of available facilities within the lower category strata<sup>2</sup>, facilities denying access for data collection, and the potential for the facility category to be redefined based on more current OSCAR data this strict facility category numbers were not maintained. However, the final sample yielded a general distribution of low versus high staffed facilities as originally designed (i.e., Category One n=8; Category Two n=24; Category Three n=40; and Category Four n=35). The investigators attempted to acquire data on more low staffed facilities, but were not able to due to facility refusals. An ongoing dialog with the government project officer was maintained during these strata adjustments to ensure the integrity of the final facility sample was maintained, especially in terms of its application in resident outcomes analyses.

### **8.2.2 Online Survey and Certification Reporting System (OSCAR) Data**

The Health Care Financing Administration's Online Survey Certification and Reporting System (OSCAR) database contains information on every nursing home in the United States that is certified by Medicare and/or Medicaid. The data are collected by the state survey and certification agencies at the time of the facility's survey (performed at least annually). The survey form instructs the facility to calculate the number of staff hours worked in the last 14 days<sup>3</sup>. Full time status is defined as 35 or more hours worked per week; part time status is less than 35 hours per week. Contract staff includes individuals and organizations under contract. The OSCAR data are based on data that are self-reported by facilities and input with minimum edit checks.

Typically, facilities are surveyed annually, as recertification must occur no less often than every fifteen months. Some facilities are surveyed more than once in a given year if there are substantial changes in a facility's organization or because of complaints about the quality of care. OSCAR calendar year files contain all facility surveys performed during that year. The beginning date of the facility's survey determines the calendar year into which the facility survey data will fall.

OSCAR staffing variables are reported in terms of FTE equivalents based on a 35 hour work week over a two week period. The conversion from FTEs to staff-hours-per-resident-day was made by

---

<sup>2</sup> In 1998, there were only 15 Ohio facilities that had fewer than 2.0 total nursing hours per resident day (excluding Directors of Nursing), after applying the OSCAR decision rules developed in this chapter.

<sup>3</sup> Note that there is some ambiguity about the time period to be recorded in OSCAR. The instructions call for facilities to use a two week period to calculate hours worked, but facilities that do not use a two week payroll period may record hours corresponding to the payroll period rather than a two week period. This is a potential source of error in the OSCAR data.

summing staff types within each staffing category (e.g., LPN hours per resident day were calculated as the sum of full time LPN full-time equivalents (FTEs), part time LPN FTEs and contracted LPN FTEs) per day for the period covered by OSCAR. Total nurse staff hours per resident per day was calculated by dividing the total staff hours per day by the average daily number of residents recorded in OSCAR. There is some ambiguity about how the number of residents is recorded in OSCAR— for example, facilities may differ with respect to how they report residents that were hospitalized during the period covered by OSCAR.

Two analytic samples were used for this study. OSCAR data from 1995-1999 were used in the study. Cleaned data for calendar year files 1995 through 1997 for Ohio, New York, and Texas were provided by Mick Cowles, of Cowles Research Group. The cleaning process involved deleting duplicate records, back-filling clearly erroneous data fields through a series of logic edits, and retaining only nursing home and skilled nursing facility observations. OSCAR data from 1998-1999 were extracted from the HCFA Data Center for facilities in the payroll sample.

### **8.2.3 Medicaid Cost Report Data**

Medicaid Cost Report data were obtained directly from New York, Texas, and Ohio for the years 1995, 1996 and 1997. Cost report data provide a comprehensive listing of facility staffing and cost variables, including contract staff, and are used by the state for facility reimbursement and accounting purposes.

Facilities report their costs annually to their state reimbursement agency, and states may use penalties against facilities misreporting data. Because the cost reports are desk audited and associated with facility reimbursement and because there are punitive measures associated with misrepresenting information, the Medicaid cost report data are considered quite accurate.

Most of the New York facilities report their Medicaid costs in calendar year format (i.e., 1/1 - 12/31). A small number of facilities, including veterans' homes, report Medicaid costs on a fiscal year starting April 1 or July 1. The Medicaid cost report data from New York includes free-standing and hospital based facilities (except a few private pay only facilities) files for 1995, 1996 and 1997. Since conversion of fiscal year data to calendar year data required extensive programming and data set resources with limited benefits in terms of averaging values across data set years to create a calendar year value, facilities reporting in fiscal year format were retained in the database with no adjustments.

Ohio facilities submit Medicaid cost report data according to calendar year. Most Texas Medicaid cost reporting periods are reported in terms of a calendar year, with a few reporting in terms of a fiscal year. Again, adjustments from fiscal year to calendar year were not made due to extensive practical

costs with limited value added in terms of analyses.<sup>4</sup>

### 8.2.4 Creation of Analytic Staffing Measure Variables

The unit of staffing measure in each data set was converted to nurse staff hours per resident day within the parameters of available data. Some decision rules were applied before final staffing variables could be derived. For example, the OSCAR data are reported in terms of FTE equivalents based on a 35 hour work week over a two week period. The conversion from FTEs to staff-hours-per-resident-day was made by summing staff types within each staffing category (e.g., LPN hours per resident day were calculated as the sum of full time LPN full-time equivalents (FTEs), part time LPN FTEs and contracted LPN FTEs) per day for the period covered by OSCAR. Total nurse staff hours per resident per day was calculated by dividing the total staff hours per day by the average daily number of residents.

## 8.3 Methods

Several types of comparisons to the payroll data were used to assess the validity of OSCAR and the Medicaid Cost Reports:

- C Comparisons of mean staffing levels, both overall and for facilities on the low end of the staffing distribution. OSCAR was used to stratify facilities to ensure adequate representation of low-staffed facilities in the payroll sample, but the payroll data were used to identify low-staffed facilities for the validity analyses discussed in this report.
- C Analysis of the proportion of facilities that had 'more valid' data in OSCAR than in the payroll data. This was a facility-level analysis that measured the proportion of the sample that had more valid staffing measures (defined by the difference between reported staffing measures from OSCAR or the Medicaid Cost Report and comparable measures from the payroll data) in OSCAR.
- C Correlation analysis. Correlation coefficients are a measure of the strength and direction of the linear relationship between two variables. The correlation between staffing measures was estimated from the three data sources using three types of correlation measures:
  - Pearson correlation coefficient: The Pearson correlation coefficient is calculated as the square root of the R-squared obtained by regressing one variable on the other. A coefficient of one indicates a linear relationship between the two variables, while a correlation coefficient of zero indicates that no relationship between the two variables is

---

<sup>4</sup> Additional information about Medicaid data and variables employed in this analysis can be found in Appendix E.

present.

--Spearman rank correlation: The Spearman correlation coefficient is the correlation of the *ranks* of the variables. Because the Pearson correlation coefficients may be greatly affected by outliers, which contribute disproportionately to the total variance of reported staffing measures, the Spearman correlations are a useful complement to the more commonly used Pearson correlation coefficients.

--Kendall's Tau-b: Kendall's Tau-b is a measure of association between ordinal variables. It is based on the number of concordant and discordant pairs of observations and uses a correction for tied pairs. The weakness of Tau-b is that it is difficult to interpret as a measure of association (or reduction in error of prediction).

## 8.4 Results

### 8.4.1 Analysis of Average Staffing Levels

Seventy eight facilities had observations from all three data sources (including both periods of payroll data). For these facilities, the mean staffing levels from both the Medicaid Cost Report data and OSCAR were similar to those derived from the Ohio payroll data (Table 1). While the differences were small, mean staffing levels from the Medicaid Cost Report data were closer to the payroll data (both periods) than were staffing measures from OSCAR.

- C Mean total nursing hours per resident day were essentially the same across all three data sources. Total hours were 3.35 in OSCAR, 3.44 in the Cost Report data and 3.43 and 3.40 in the two payroll periods (Table 8.1).
- C Mean RN and LPN hours per resident day were also basically identical across the three data sources. For the pre-survey period, mean RN and LPN hours were 1.21 in both OSCAR and the payroll data (survey period), compared to 1.28 in the Cost Report data. In the pre-survey period, mean RN and LPN hours in the payroll data were 1.28.

Low-staffed facilities were emphasized in the validity analyses. Among facilities that ranked in the bottom 20 in terms of total hours per resident day (based on the survey period), mean staffing measures from OSCAR were somewhat closer to the payroll data than were the Medicaid Cost Report data.

- C For low staffed facilities, reported staffing levels were higher in the Cost Report data (an average of 2.99 nursing hours per resident day) and OSCAR (an average of 2.73 nursing hours per resident day) than in the payroll data, which had a mean of 2.65 hours per resident day in the pre-OSCAR period and 2.31 in the OSCAR period.



- C The same general pattern was also observed with respect to the sum of RN and LPN hours per resident day. Mean RN and LPN hours were 1.09 in the Cost Report data, compared to 0.99 in OSCAR, and 0.84 and 0.88 in the two payroll periods.

It is important to note that the similarities in mean staffing levels do not imply that staffing measures for individual facilities across the three facilities are highly correlated. As the correlation results discussed in the next section show, there were some facilities that had large differences in staffing values for individual facilities across the different data sources.

Table 8.1: Comparison of reported staffing levels from Ohio payroll data to Medicaid Cost Report and OSCAR– Average staffing levels				
	Ohio payroll data		Medicaid Cost Report	OSCAR
	pre-survey period	survey period		
All facilities				
Total hours per resident day	3.43 (1.28)	3.40 (1.55)	3.44 (0.96)	3.35 (1.31)
RN and LPN hours per resident day	1.21 (0.79)	1.29 (0.98)	1.28 (0.65)	1.21 (0.68)
Low staffed facilities- Bottom 20 in total weighted nursing hours per resident day				
Total hours per resident day	2.65 (0.70)	2.31 (0.46)	2.99 (0.46)	2.73 (0.51)
RN and LPN hours per resident day	0.88 (0.27)	0.84 (0.33)	1.09 (0.26)	0.99 (0.39)
N= 78 Standard deviations shown in parentheses. Sources: Ohio payroll data, OSCAR, Medicaid Cost Report				

## 8.4.2 Correlation Analysis

Staffing measures from the Medicaid Cost Reports were more highly correlated with payroll data than were the OSCAR-based measures (Tables 8.2 and 8.3). Based on the correlation analyses, the Medicaid Cost Report data are the more valid source of staffing data, especially considering that the payroll data were collected for the same period as OSCAR, while the Cost Report data were based on average figures across an entire year. Comparison of staffing levels from the two periods included in the payroll data suggested that there was considerable fluctuation in staffing levels across time (see discussion in Chapter 7): the Pearson correlation for total nursing hours per resident day between the two payroll periods is .76.

For the pre-survey period, most of the correlation analyses suggested that the Cost Report data were more consistent with the payroll data than were staffing measures from OSCAR:

- C For the pre-survey period, the Pearson correlation coefficient for total hours per resident day was higher between OSCAR and the payroll data (0.761) than between the Cost Report and payroll data (0.697) (Table 8.2). The correlation for total RN and LPN hours was slightly higher for OSCAR (0.811) than for the Cost Report data (0.794).
- C The Spearman rank correlation for total hours was higher for the Cost Report data (0.610) than for OSCAR (0.532), indicating that the higher Pearson correlation for OSCAR was due to its consistency for facilities with very low or very high staffing levels. The Kendall Tau-B statistic was 0.473 for the Cost Report, compared to 0.406 for OSCAR. For RN and LPN hours, both the Spearman correlation coefficient and the Kendall Tau-B statistic were about 0.200 higher for OSCAR than the payroll data.
- C Excluding the ten facilities with the highest staffing levels, all three correlation measures were considerably higher for the Cost Report data than for OSCAR. For total hours per resident day, the Pearson correlation coefficient was 0.611 for the Cost Report, compared to 0.454 for OSCAR, and the Spearman correlation coefficient was 0.641 for the Cost Report, higher than the 0.538 observed for OSCAR. For RN and LPN hours, all three correlation statistics were higher for the Cost Report data than for OSCAR. The Pearson correlation coefficient was 0.554 for the Cost Report data, much higher than for OSCAR (0.331).
- C Among lower staffed facilities (those with below-median total hours reported in the payroll data), the Cost Report data were more highly correlated with the payroll data than OSCAR. For these facilities, the Pearson correlation in total hours was 0.559 for the Cost Report data, compared to

0.459 for OSCAR. Both the Spearman and Kendall correlations were also higher for staffing measures based on Cost Report data. For RN and LPN hours, the Pearson correlation coefficient was much higher for the Cost Report data (0.706) than for OSCAR (0.252). The greater consistency of the payroll and Cost Report data for low-staffed facilities was an important criterion in evaluating the reliability of the Cost Report data and OSCAR.

Even for payroll records covering the same time period as the OSCAR assessment, analyses suggested that the Cost Report data were more valid and reliable than OSCAR. Oddly, the correlation between the payroll data and OSCAR was consistently lower for the survey period than the preceding period (as were the correlations between the Cost Report and payroll data).

- C For the survey period, the Pearson correlation coefficient for total hours was 0.594 for the Cost Report data but only 0.490 using OSCAR (Table 8.3). The Spearman and Kendall correlations were also both higher for the Cost Report data than for OSCAR. For RN and LPN hours, this correlation was 0.687 for the Cost Report data and only 0.521 in OSCAR. The Spearman and Kendall Tau-B correlation coefficients were both also higher for OSCAR than for the Cost Report data.
- C Excluding the ten highest staffed facilities (defined using the payroll data), for total hours, all of the correlation statistics were considerably higher for the Cost Report data. The Pearson correlation was 0.444 for the Cost Report data, compared to 0.265 for OSCAR, and the Spearman correlation was 0.445 for the Cost Report data and 0.391 in OSCAR. For combined RN and LPN hours, the Pearson correlation coefficient was higher for the Cost Report data, but both the Spearman correlation and Kendall's Tau-B were higher for OSCAR.
- C Among lower staffed facilities (those with below median total nursing hours), the payroll data were more highly correlated with the Cost Report data than for OSCAR. The Pearson correlation coefficient for total hours was 0.310 for the Cost Report data, much higher than the 0.121 for OSCAR. Spearman and Kendall Tau-B correlation coefficients were both also considerably higher for the Cost Report data. The Kendall Tau-B statistic was 0.236 for the Cost Report data and only 0.126 for OSCAR. For these facilities, all three correlation measures for combined RN and LPN hours were higher for the Cost Report data than for OSCAR.

Based on the correlation analyses, the Cost Report data are the more valid and reliable data source. Most of the correlation coefficients were higher for the Cost Report data, especially those for facilities with lower staffing levels. Given OSCAR's advantage in covering the same time period as one of the payroll periods, the results even more strongly suggest the superior reliability of the Cost Report data.

There are legitimate reasons for differences between the payroll and Cost Report data, given the different time periods that they cover, but differences in staffing levels from OSCAR and the payroll survey period are almost certainly due to inaccuracies in the OSCAR data.

<b>Table 8.2: Consistency of reported staffing levels from Ohio payroll data to Medicaid Cost Report and OSCAR: Correlation analysis– pre-survey period</b>						
	<b>Correlation with payroll data–pre-survey period</b>					
	<b>Pearson correlation</b>		<b>Spearman rank correlation</b>		<b>Kendall's Tau-b</b>	
<b>Staffing measure</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>
<b>All included facilities</b>						
Total hours per resident day	0.697	0.761	0.610	0.532	0.473	0.406
RN and LPN hours per resident day	0.794	0.811	0.565	0.585	0.432	0.453
<b>Excluding top 10 facilities (based on payroll data, total staffing)</b>						
Total hours per resident day	0.611	0.454	0.641	0.538	0.481	0.390
RN and LPN hours per resident day	0.554	0.331	0.539	0.520	0.404	0.386
<b>Low staffed facilities-- lowest 50% (based on payroll data, total staffing)</b>						
Total hours per resident day	0.559	0.459	0.625	0.579	0.465	0.403
RN and LPN hours per resident day	0.706	0.252	0.609	0.445	0.487	0.311
N= 78 Sources: Ohio payroll data, OSCAR, Medicaid Cost Report						

<b>Table 8.3: Consistency of reported staffing levels from Ohio payroll data to Medicaid Cost Report and OSCAR: Correlation analysis– Survey period</b>						
<b>Staffing measure</b>	<b>Correlation with payroll data–survey period</b>					
	<b>Pearson correlation</b>		<b>Spearman rank correlation</b>		<b>Kendall’s Tau-b</b>	
	<b>Medicaid Cost Report</b>	<b>OSCAR</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>
<b>All included facilities</b>						
Total hours per resident day	0.594	0.490	0.449	0.392	0.343	0.297
RN and LPN hours per resident day	0.687	0.521	0.509	0.610	0.391	0.486
<b>Excluding top 10 facilities (based on payroll data, total staffing)</b>						
Total hours per resident day	0.444	0.265	0.445	0.391	0.337	0.274
RN and LPN hours per resident day	0.602	0.455	0.540	0.629	0.413	0.473
<b>Low staffed facilities-- lowest 50% (based on payroll data, total staffing)</b>						
Total hours per resident day	0.310	0.121	0.332	0.196	0.236	0.126
RN and LPN hours per resident day	0.563	0.269	0.551	0.471	0.436	0.344
N=78 <i>Sources:</i> Ohio payroll data, OSCAR, Medicaid Cost Report						

### 8.4.3 Consistency of Identification of Low-Staffed Facilities

An important criterion in the validity analyses is the ability of staffing measures to correctly identify low-staffed facilities. The distribution of low-staffed facilities was measured (defined based on the payroll data using a variety of thresholds) in terms of staffing levels from OSCAR and the Cost Reports. The Cost Report data more accurately identified low-staffed facilities, although most of the differences were small.

- C In the pre-survey period, 5 of the 10 lowest-staffed facilities (based on the payroll data) were among the ten lowest staffed facilities according to the Cost Report data. OSCAR correctly placed two of these facilities among the bottom 10 (Table 8.4). The Cost Report data were also more accurate in identifying the ten lowest staffed facilities from the survey period.
- C The performance of OSCAR and the Cost Report data in identifying the 20 lowest-staffed facilities was virtually identical for both periods. For both data sources, 50% of facilities in the survey period were among the 20 lowest staffed facilities. In the pre-survey period, 60% of these facilities were among the 20 lowest staffed facilities in the Cost Report data, compared to 55% in OSCAR.
- C The ability of OSCAR and the Cost Report data to identify the 30 and 40 lowest staffed facilities was quite similar, for both the pre-survey and survey periods.

The investigators also analyzed the distribution of Cost Report and OSCAR staffing measures stratified by staffing levels from the payroll data. Among the 20 lowest staffed facilities in the survey period, both OSCAR and the Cost Report data also had these facilities as one of the 20 lowest staffed facilities. In the pre-survey period, 60% of these facilities were among the 20 lowest staffed facilities in the Cost Report data, compared to 55% in OSCAR. Both the Cost Report data and OSCAR ranked 90% of these facilities as one among the 40 lowest staffed facilities.

- C The ability of the two data sources to correctly identify low-staffed facilities was comparable, although the Medicaid Cost Report data performed somewhat better, especially for facilities with very low staffing levels. While no firm conclusions can be drawn based on the analyses presented in Tables 8.4 - 8.6, these analyses do not change the conclusion, based on the correlation analyses presented in Tables 8.2 and 8.3 that the Cost Report data are more valid and reliable than OSCAR. It is important to note that the payroll sample did not include many low staffed facilities, as the proportion of facilities with fewer than 2.0 nursing hours per resident day was much lower in Ohio than for the nation as a whole, and, as a result, the payroll sample

did not include as many low-staffed facilities as desired.

<b>Table 8.4:</b> <b>Analysis of how consistently OSCAR and the Medicaid Cost Report data identify low-staffed facilities</b>				
<b>Subset of facilities</b>	<b>Proportion of facilities correctly identified</b>			
	<b>Period preceding OSCAR assessment</b>		<b>OSCAR assessment period</b>	
	<b>Medicaid Cost Report</b>	<b>OSCAR</b>	<b>Medicaid Cost Report</b>	<b>OSCAR</b>
Bottom 10 (based on payroll data)	50%	20%	30%	10%
Bottom 15 (based on payroll data)	53%	33%	47%	20%
Bottom 20 (based on payroll data)	60%	55%	50%	50%
Bottom 30 (based on payroll data)	63%	67%	60%	63%
Bottom 40 (based on payroll data)	73%	68%	72%	70%
N= 78 <i>Sources:</i> Ohio payroll data, OSCAR, Medicaid Cost Report				

<b>Table 8.5:</b> <b>Analysis of ability of Medicaid Cost Report data to identify low-staffed facilities</b>				
<b>Distribution of facilities</b>	<b>Distribution of facilities in Medicaid Cost Report data</b>			
	<b>Bottom 10</b>	<b>Bottom 20</b>	<b>Bottom 30</b>	<b>Bottom 40</b>
<b>Pre-survey period</b>				
Bottom 10 (based on payroll data)	50%	90%	90%	90%
Bottom 20 (based on payroll data)	35%	60%	80%	90%
Bottom 30 (based on payroll data)	23%	40%	63%	80%
<b>Survey period</b>				
Bottom 10 (based on payroll data)	30%	60%	60%	60%
Bottom 20 (based on payroll data)	30%	50%	65%	70%
Bottom 30 (based on payroll data)	23%	40%	60%	73%
N= 78 <i>Sources:</i> Ohio payroll data, OSCAR, Medicaid Cost Report				



<b>Table 8.6:</b> <b>Analysis of ability of OSCAR to identify low-staffed facilities</b>				
<b>Distribution of facilities</b>	<b>Distribution of facilities in OSCAR</b>			
	<b>Bottom 10</b>	<b>Bottom 20</b>	<b>Bottom 30</b>	<b>Bottom 40</b>
<b>Pre-survey period</b>				
Bottom 10 (based on payroll data)	20%	70%	80%	90%
Bottom 20 (based on payroll data)	10%	55%	80%	90%
Bottom 30 (based on payroll data)	10%	43%	63%	80%
<b>Survey period</b>				
Bottom 10 (based on payroll data)	10%	50%	60%	70%
Bottom 20 (based on payroll data)	15%	50%	70%	75%
Bottom 30 (based on payroll data)	10%	40%	63%	77%
N= 78 <i>Sources:</i> Ohio payroll data, OSCAR, Medicaid Cost Report				

#### **8.4.4 Conclusion of Validity Analyses: Medicaid Cost Report Data Are More Valid Than OSCAR**

Staffing measures from the Cost Report data are more valid than those of OSCAR, and the Cost Report should be the source of staffing measures for analyses of the relationship between staffing and resident outcomes. Most of the correlation measures were higher for the Cost Report data, especially for low staffed facilities. Even based on a comparison to OSCAR data that covered the same period of time as the payroll data, total hours from the Cost Report data were more highly correlated with the payroll data than were OSCAR data.

The investigators placed emphasis on the validity of OSCAR and Medicaid Cost Report data for facilities at the low end of the staffing distribution. With a total sample of 78 facilities for which data from all three sources (including both periods of payroll data) were available, there were relatively few low-staffed facilities on which to base these analyses. Correlation coefficients for facilities with below-median staffing levels were much higher for Medicaid Cost Report-based measures than for OSCAR. The Medicaid Cost Reports also had slightly better ability to correctly identify low-staffed facilities.

While the investigators recommend that the Cost Report data be used for the outcomes study, there were facilities for which the Cost Report data appeared to be unreliable. The decision rules discussed

in Section 8.5 can be used to determine which facilities to exclude from the analysis sample, resulting in a subset of facilities with valid and reliable OSCAR data that can be used for the outcomes study.

## **8.5 Identifying Unreliable Staffing Data**

It is important to develop a set of exclusion criterion for identifying facilities with unreliable staffing measures in the Cost Report data. Because of the small number of facilities for which payroll data were available, these were developed independently of the payroll data.

Two types of potential decision rules were developed to exclude facilities with questionable staffing measures. The first is a set of “logical” decision rules that identified a relatively small subset of facilities for which reported staffing levels did not make sense. These are similar to the OSCAR decision rules described in Chapter 7, although the Cost Report data did not allow us to implement all of these rules.

The second set of decision rules can be used to identify facilities that had large changes in reported staffing levels across time that could not be validated using OSCAR. Staffing measures for facilities that report extremely large changes in staffing levels from one year to the next are questionable, although these changes can occur. It may be desirable to exclude facilities that report large changes if these changes cannot be validated using OSCAR (or some other independent data source). A limitation of this type of exclusion criteria is that they rely on OSCAR to validate changes in Cost Report data, despite the fact that the Cost Report data were found to be more valid and reliable than OSCAR.

A general principle guiding the set of decision rules that was applied for developing a file containing valid and reliable staffing data were that, except in the case of extreme outliers that represent obvious data errors, no facilities were excluded simply because of the facility’s *level* of staffing. Rather, exclusion decisions were based on the *reliability* of the facility’s staffing measures, based on across-time consistency and the consistency of staffing measures from the Cost Report and OSCAR data.

### **8.5.1 Identification of Extreme Outliers**

There were fifteen facilities in the three state sample for which the Cost Report data reported less than 0.5 total hours per resident day, including two with zero hours reported. Six facilities were reported with between 12 and 7,714 total hours per resident day. While the number of facilities with these type of obvious data errors is small, they contributed disproportionately to the total across-time variance in the Medicaid Cost Report data.

The investigators were reluctant to exclude the low-staffing outliers, as legitimately low staffed facilities

were of special interest for the outcomes analysis. Four facilities reported less than 0.5 total hours per resident day for both 1996 and 1997 in the Medicaid cost report data, but none of these facilities appeared to have low staffing levels based on OSCAR. For the other facilities reporting fewer than 0.5 total nursing hours in the Cost Report data, staffing levels for the other year (either 1996 or 1997) were 2.2 or higher (This was within one standard deviation of mean total hours, which were around 3.0 for the three state sample). This suggested that the low-staffing outlier values did not reflect actual staffing at the facility but rather data errors. The investigators recommend excluding all facilities that report fewer than 0.5 total nursing hours per resident day.

The investigators evaluated whether a higher threshold should be used to exclude low-cost outliers, but, given our interest in preserving the distribution of staffing levels across the sample decided not to automatically exclude other facilities on the basis of their reported staffing levels. It should be noted, however, that most of the facilities that reported between 0.5 and 1.0 hours per resident day would be excluded based on the longitudinal decision rule described below.

### **8.5.2 Exclusions Based on Changes in Staffing Levels Across Time**

Except for the extreme outliers that have obvious data errors, the reliability of individual observations requires subjective judgments regarding what threshold of change in staffing across time represents unreliable data as opposed to normal fluctuations in staffing level. All facilities were ranked on the percentile change in total hours per resident day and evaluated how the consistency of staffing measures across time changed depending on which facilities were included in the analysis.

#### *8.5.2.1 Correlation of Medicaid Cost Report Staffing Measures Across Time*

The overall correlation between total hours per resident day based on the Cost Report data was 0.806 between 1995 and 1996 and 0.774 between 1996 and 1997 (Note that facilities with fewer than 0.5 or more than 12 total nursing hours per resident day were excluded from these analyses) (Table 8.11). By construction, exclusion of facilities with changes in staffing will lead to improvement in the correlation of staffing measures across periods, and this does not necessarily mean that the reliability of data for the remaining facilities has been improved.

- C When the 1% of facilities with the largest change in staffing were excluded, the correlation increased to 0.868 for 1995-96 and to 0.889 for 1996-97.
- C Excluding the 5% of facilities with the largest change in staffing, the correlation coefficient for both 1995-96 and 1996-97 was above 0.90.

- C Excluding 10% of facilities with the largest change in across-year total staffing, the correlation for the remaining facilities was 0.93 for 1995-96 and 0.94 for 1996-97.

The correlation analyses presented in Table 8.11 show a trade-off between increasing the across-time consistency of staffing measures and decreasing the sample size available for outcomes analyses. Because the criteria used to decide which facilities to keep in the analysis were not based on a facility's actual *level* of staffing, but rather the consistency of staffing measures, the distribution of staffing that is found in the data is preserved.

#### *8.5.2.2 Correlation for Low-Staffed Facilities*

Given the research questions to be addressed by the outcomes study, it is important that the analysis sample include valid and reliable data for low-staffed facilities. The across-time correlation was evaluated for facilities in the lowest 15% in terms of total nursing hours per resident day.

Among these facilities, the across-time correlations were lower than for the full sample, but, for 1996-1997, a correlation above 0.71 can be obtained if the 15% of facilities with the largest change in staffing are excluded (Table 8.12). It is important to note that this relatively high correlation does not prove the accuracy of staffing measures for the remaining facilities, only that their staffing measures are consistent across time.

<b>Table 8.11: Correlation between staffing measures across time based on different exclusion criterion</b>		
<b>Subset of facilities included</b>	<b>Pearson correlation coefficient– Total nursing hours per resident day</b>	
	<b>1995-1996 (N)</b>	<b>1996-1997 (N)</b>
Exclude facilities that failed one or more of the logical thresholds	0.806 (1,600)	0.774 (1,962)
Exclude 1% of facilities with largest percentile change (including all facilities with 100% or more change)	0.868 (1,546)	0.889 (1,919)
Exclude 5% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.902 (1,431)	0.924 (1,848)
Exclude 10% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.926 (1,406)	0.942 (1,742)
Exclude 15% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.936 (1,332)	0.952 (1,657)
<i>Sources:</i> Medicaid Cost Reports, 1995-1997		

<b>Table 8.12:</b> <b>Correlation between staffing measures across time based on different exclusion criterion– Facilities in lowest 15% of total hours per resident day only</b>		
<b>Subset of facilities included</b>	<b>Pearson correlation coefficient– Total nursing hours per resident day</b>	
	<b>1995-1996 (N)</b>	<b>1996-1997 (N)</b>
Exclude facilities that failed one or more of the logical thresholds	0.269 (265)	0.193 (304)
Exclude 1% of facilities with largest percentile change in staffing (including all facilities with 100% or more change)	0.510 (236)	0.361 (295)
Exclude 5% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.656 (228)	0.544 (287)
Exclude 10% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.677 (223)	0.713 (274)
Exclude 15% of facilities with largest percentile change in staffing or a 100% or more change in total hours per resident day	0.703 (215)	0.725 (266)
<i>Sources:</i> Medicaid Cost Reports, 1995-1997		

## 8.6 Potential Exclusion Criteria for Medicaid Cost Report Staffing Measures

Other than for a relatively small number of facilities with obvious data errors, however, there is no single definition of what the characteristics of valid and reliable data should be. The previous section showed the trade-offs between preserving the sample and increasing the across-time correlation in Cost Report staffing measures

One type of decision rule that could be implemented is to exclude facilities in the top 10% in terms of across-time change in total nursing hours unless this change can be validated based on OSCAR (i.e., whether the facility was also reported in the top 10% in terms of across-time change in total hours). Such an approach allows data for facilities with large changes to remain in the sample if a similar change is observed in OSCAR.

This type of decision rule improved the correlation of staffing measures from OSCAR and the Cost Report data:

- C Implementation of this decision rule improved the Pearson correlation between 1997 total hours from the Cost Report data and OSCAR from 0.480 (using only the logical decision rules for OSCAR and the Cost Report data) to 0.550 (Table 8.13). The Spearman correlation coefficient and Kendall's Tau-B also increased after application of this decision rule.
- C Excluding facilities with large changes in total hours that could not be validated using OSCAR also increased the correlation of RN and LPN hours from the Cost Report and OSCAR data, although the increase was smaller than for total hours. The Pearson correlation increased from 0.552 to 0.567 using this decision rule.

While use of OSCAR data to validate changes in the Cost Report data improves the consistency of staffing measures from the two data sources, one problem with this approach is that it uses a less valid data source (OSCAR) to validate data from a more reliable source (the Cost Reports). While the investigators found that the Cost Report data were more reliable than OSCAR, the differences were small, and it seems more likely that large changes in staffing levels are more likely to be valid if they can be validated using a second, independently collected data source such as OSCAR, in spite of potential reliability concerns.

The decision about whether to apply decision rules based on the across time change in staffing levels is one that ultimately must be made by the analyst, based on the trade-off between maintaining an adequately sized sample and increasing the reliability of staffing measures for facilities that are included in the analyses.

<b>Table 8.13: Impact of Medicaid Cost Report Decision Rules on Consistency with OSCAR Data: Correlation analysis</b>			
<b>Type of decision rule used</b>	<b>Correlation with payroll data</b>		
	<b>Pearson correlation</b>	<b>Spearman rank correlation</b>	<b>Kendall's Tau-B</b>
<b>Logical decision rules only</b>			
Total hours per resident day	0.480	0.601	0.439
RN and LPN hours per resident day	0.552	0.678	0.501
<b>Logical decision rules and decision rule based on change in total hours</b>			
Total hours per resident day	0.550	0.633	0.465
RN and LPN hours per resident day	0.567	0.704	0.524
N= 1,474 facilities with 1997 OSCAR and Medicaid Cost Report data <i>Sources:</i> OSCAR, Medicaid Cost Report			

## 8.7 Conclusions

Based on comparisons to Ohio payroll data, staffing measures from the Cost Report data were more valid than those of OSCAR. Even based on a comparison to OSCAR data that covered the same period of time as the payroll data, total hours from the Cost Report data were more highly correlated with the payroll data than were OSCAR data. Most of the correlation measures were higher for the Cost Report data, especially for low staffed facilities.

Even if the Cost Report data were totally accurate, we would not expect a perfect (i.e., 1.0) correlation between the Cost Report staffing measures representing an average for the entire year and the payroll (“gold-standard”) measure for a 2-week period within that year. Under these circumstances, the correlation between these two accurate measures would depend on the degree of variability in staffing levels throughout the year. And, as we have seen, there is considerable fluctuation - the correlation between the two “gold-standard” payroll periods was .76, a strong correlation to be sure, but one permitting a considerable amount of fluctuation over the course of a year. Hence, the .59 correlation for total nursing hours between the Cost Report data and the payroll data (survey period) is actually a closer association than the correlation would suggest. For all these reasons, we conclude that the Cost Report data are not only better than OSCAR, but may be quite good. Hence, the Cost Reports should



be the source for the staffing measures used in the outcomes analyses.

At a minimum, the small number of facilities that report extreme outlier staffing values (total nursing hours per resident day less than 0.5 or greater than 12) should definitely be excluded from the analytic file. The consistency of staffing measures between the Cost Reports and OSCAR can also be improved by excluding facilities that reported large change in staffing levels across time that could not be validated using OSCAR. The decision about whether to apply this type of decision rule ultimately rests with the data analyst, as it depends on the tradeoff between having a large sample versus having a staffing file that likely has somewhat greater reliability.